RANDOM AUTOCATALYTIC NETWORKS

E MOSSEL and M Steel

Department of Mathematics and Statistics
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

Report Number: UCDMS2004/15 JUNE 2004

RANDOM AUTOCATALYTIC NETWORKS

ELCHANAN MOSSEL, MIKE STEEL

ABSTRACT. We determine conditions under which a random biochemical system is likely to contain a subsystem that is both autocatalytic and able to survive on some ambient 'food' source. Such systems have previously been investigated for their relevance to origin-of-life models. In this paper we extend earlier work, by finding precisely the order of catalysation required for the emergence of such self-sustaining autocatalytic networks. This answers questions raised in earlier papers, yet also allows for a more general class of models. We also show that a recently-described polynomial time algorithm for determining whether a catalytic reaction system contains an autocatalytic, self-sustaining subsystem is unlikely to adapt to allow inhibitory catalysation - in this case we show that the associated decision problem is NP-complete.

Address for correspondence:

Mike Steel

Biomathematics Research Centre, University of Canterbury

Private Bag 4800, Christchurch, New Zealand

Email: m.steel@math.canterbury.ac.nz

Fax: 64-3-3642587.

Date: 14 May 2004.
Key words and phrases. combinatorial chemistry, autocatalysis, discrete random structures.

1. Introduction

The idea that the study of discrete random networks could provide some insight into the problem of how primitive life might have emerged from an ambient 'soup' of molecules goes back the mid-1980s. This was largely motivated by the earlier investigation of random graphs, pioneered by Alfred Rényi and Paul Erdös in the 1950's and 1960s, which had revealed the widespread occurrence of 'threshold phenomena' (sometimes also called 'phase transitions') in properties of these graphs. In the simplest random graph model one has set of vertices (points) and edges are added independently and randomly between pairs of vertices. As the probability that any two nodes are jointed by an edge passes certain well-studied thresholds, there is typically a fundamental change in various qualitative properties of a large random graph, such as its connectivity, or the size of the largest component (see eg. Bollobas, 2001). Extending this approach, Bollobas and Rasmussen (1989) investigated when a directed cycle would first emerge in a random directed graph, and how many vertices such a cycle would contains. They were motivated by the idea that the emergence of a primitive metabolic cycle was an essential step in the early history of life, writing "we want to know when the first catalytic feedbacks appear, and how many different RNA molecules they involve." Cohen (1988) also foresaw the relevance of random graph techniques for modelling primitive biological processes.

The importance of cycles in early life had also been studied - from a slightly different perspective - by Eigen (1971) and Eigen and Schuster (1979). They proposed a metabolic 'hypercycle' as a way of circumventing the so-called 'error catastrophe' in the formation of longer strings of nucleotides, first demonstrated by Maynard-Smith (1983). The study of such processes and how they might further evolve into early life has been extensively investigated, using both stochastic and dynamical approaches (eg. Scheuring, 2000; Wills and Henderson, 1997; Zintzaras, Santos and Szathmáry, 2002).

The idea that threshold phenomena might help explain some of the mystery surrounding the emergence of life-like systems from a soup of inanimate molecules was developed further by Stuart Kauffman (1986, 1993). He realised simple random graphs and digraphs by themselves do not capture the intricacy of chemical reactions and catalysis. A more complex discrete structure - which has become known as a *catalytic reaction system* is required in order to formalize and study the concept of a system of molecules that catalyses all the reactions required for their generation, and which can be sustained from some ambient 'food' source of molecules, F. Several investigators have developed this type of approach (Hordijk and Fontanari, 2003; Hordijk and Steel, 2004; Lohn *et al.*, 1998; Wills and Henderson, 1997) though it also has its critics (eg. Lifson, 1997).

At this point there are two possible ways to formalize the concept of a self-sustaining and autocatalytic set of molecules - these are referred to here as the RAF (reflectively autocatalytic, and F–generated) and CAF (constructively autocatalytic and F–generated) sets. The former was investigated in Steel (2000) and Hordijk and Steel (2004). A CAF, which we formalize in this paper is a slightly

stronger notion - it requires that any molecule m that is involved in any catalysation must already have been built up from catalysed reactions (starting from F). This concept is perhaps overly restrictive, since it might be expected that m would still be present in a random biochemical system in low concentrations initially before reactions that generate a steady supply of m become established. Thus although we present results for CAFs, our primary interest is in RAFs.

For the sequence-based models studied by Kauffman, we determine the degree of catalysation required for a RAF or a CAF to arise. In Kauffman's model reactions consist of the concatenation and cutting of sequences up to some maximal (large) length, starting from small sequences of length at most t, and each molecule has a certain probability of (independently) catalysing any given reaction. Let $\mu(x)$ denote the average number of (concatenation) reactions that sequence x catalyses, which may depend on $|x|$ the length of x. Then, roughly speaking, our results show that if $\mu(x)/|x|$ is small the probability that the system contains a RAF is small; conversely if $\mu(x)/|x|$ is large the probability the system contains a RAF is close to 1, and indeed in this case there is likely to be a RAF for which all the molecules in the system are involved. This confirms two conjectures that were posed in Steel (2004) and confirms some trends that were suggested by simulations in Hordijk and Steel (2004).

Our results for RAFs contrast sharply with the degree of catalysation required for a CAF. In that case each molecule needs to catalyse, on average, some fixed proportion of all reactions for a CAF to be likely. That is, the corresponding value of $\mu(x)$ required for a likely occurrence of a CAF is exponentially larger (with n) than for a RAF.

We begin this paper by formalizing the concepts of RAF and CAF, and we do so in a more general setting than Hordijk and Steel (2004) as we consider the effect of general catalysation regimes - for example by allowing certain molecules to inhibit certain reactions. In this case determining whether an arbitrary catalytic reaction system contains a RAF seems to be computationally intractable. Indeed in Section 3 we show that the decision problem is NP-complete. This contrasts with the situation where one allows only positive catalysation; in that case a polynomial-time algorithm (in the size of the system) for finding a RAF if one exists was described in Hordijk and Steel (2004). Sections 4 and 5 present the main results concerning the required growth of $\mu(x)$ with $|x|$ required for RAF and CAR generation, and in Section 6 we make some final brief comments.

2. Preliminaries and definitions

We mostly follow the notation of Steel (2000) and Hordijk and Steel (2004). Let X denote a set of molecules and $\mathcal{R}$ a set of reactions, where we regard a reaction as an ordered pairs (A, B) where A, B are subsets of X called the *reactants* and *products* respectively. Let F be a distinguished subset of X, which can be regarded as some plentiful supply ('food') of reactants.

For $r \in \mathcal{R}$ let $\rho(r) = A$ and $\pi(r) = B$ and for a set $\mathcal{R}' \subseteq \mathcal{R}$ let

$$\rho(\mathcal{R}') := \cup_{r\in\mathcal{R}'}\rho(r),$$

$$\pi(\mathcal{R}') := \cup_{r\in\mathcal{R}'}\pi(r),$$

and

$$\text{supp}(\mathcal{R}') := \rho(\mathcal{R}') \cup \pi(\mathcal{R}').$$

Thus $\text{supp}(\mathcal{R}')$ denotes the molecules in X that are used or produced by at least one reaction in $\mathcal{R}'$.

Given a subset $\mathcal{R}'$ of $\mathcal{R}$ and a subset X' of X the *closure* of X' relative to $\mathcal{R}'$, denoted $cl_{\mathcal{R}'}(X')$ is the (unique) minimal subset W of X that contains X' and that satisfies the following condition for each reaction $(A, B) \in \mathcal{R}'$:

$$A \subseteq X' \cup W \Rightarrow B \subseteq W.$$

It is easily seen that $cl_{\mathcal{R}'}(X')$ is precisely the set of molecules that can be generated starting from X' and repeatedly applying reactions selected (only) from $\mathcal{R}'$.

Let $c : 2^X \times \mathcal{R} \to \{0, 1\}$ be a *catalysation function*. The function c tells us whether or not each reaction r can proceed in its environment (eg. be 'catalysed') depending on what other molecules are present. Thus let $c(A, r) = 1$ precisely when r would be catalyzed if the other molecules in the system comprise the set A. For example, consider a simple scenario where each reaction $r \in \mathcal{R}$ is catalysed provided that at least one molecule in some set (specific to r) is present. We can represent the associated function c as follows – we have a set $C \subseteq X \times R$ (as in Steel, 2000; Hordijk and Steel, 2004) where (x, r) indicates that molecule x catalyses reaction r. The catalysation function $c = c_C$ for this simple setting is then defined by

$$c_C(A, r) = \begin{cases} 1, & \text{if } \exists x \in A : (x, r) \in C; \\ 0, & \text{otherwise.} \end{cases}$$

More generally, suppose we have two sets $C(+) \subseteq X \times R$ and $C(-) \subseteq X \times R$ then a candidate for c is the function $c = c_{C(+),C(-)}$ defined by:

(1)

$$c_{C(+),C(-)}(A, r) = \begin{cases} 1, & \text{if } \exists x \in A : (x, r) \in C(+) \text{ and there is no } x' \in A : (x', r) \in C(-); \\ 0, & \text{otherwise.} \end{cases}$$

Thus $c_{C(+),C(-)}$ allows both catalysation and inhibition. We find it useful to write $A \to B$ to denote the reaction (A, B). Similarly, we will write $A \xrightarrow{C(+),C(-)} B$ to denote the reaction (A, B) together with a catalysation function that satisfies (1). When the sets A, B, C are singletons we will often omit the $\{\}$ symbols.

In case c is monotone in the first co-ordinate (i.e. $A \subset B \Rightarrow c(A, r) \leq c(B, r)$) we will call c *monotone*. Note that c_C is monotone, and that monotone catalytic functions do not allow inhibition effects.

The triple $\mathcal{Q} = (X, R, c)$ is called a *catalytic reaction system.*

2.1. **Autocatalytic networks.** Suppose we are given a catalytic reaction system $\mathcal{Q} = (X, \mathcal{R}, c)$ and a subset F of X.

A *reflexive autocatalytic network over* F or RAF for $\mathcal{Q}$ is a non-empty subset $\mathcal{R}'$ of $\mathcal{R}$ for which

(i) $\rho(\mathcal{R}') \subseteq cl_{\mathcal{R}'}(F)$
(ii) For each $r \in R', c(\text{supp}(\mathcal{R}'), r) = 1$.

In addition, to avoid biological triviality, we will also require that any RAF $\mathcal{R}'$ also satisfies the condition

(iii) $\pi(\mathcal{R}') \not\subseteq F$

Thus for $\mathcal{R}'$ to be a RAF, each molecule involved in $\mathcal{R}'$ must be able to be constructed from F by repeated applications of reactions that lie just in $\mathcal{R}'$ (condition (i)) and each reaction in $\mathcal{R}'$ must be catalysed by the system of molecules involved in $\mathcal{R}$ (condition (ii)). This definition is a slight generalization of that given by Hordijk and Steel (2004) to allow for more general catalysation functions c in condition (ii). Condition (iii) simply ensures that any set of reactions that produce only molecules that are already in the food set F does not constitute a RAF.

We also describe a condition which is somewhat stronger than the RAF requirement.

A *constructively autocatalytic network over* F or CAF for $\mathcal{Q}$ is a strictly nested sequence $\emptyset \neq \mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_k$, for which

(i) $\rho(\mathcal{R}_1) \subseteq F$ and for each $r \in \mathcal{R}_1$, $c(F, r) = 1$.
(ii) For all $i \in \{1, \ldots, k-1\}$, $\rho(\mathcal{R}_{i+1}) \subseteq \text{supp}(\mathcal{R}_i)$, and for each $r \in \mathcal{R}_{i+1}$, $c(\text{supp}(\mathcal{R}_i), r) = 1$.
(iii) $\pi(\mathcal{R}_1) \not\subseteq F$.

Informally, a CAF is a way to sequentially build up a set of molecules, starting with F, and in such a way that every reaction is catalyzed by at least one molecule that has already been constructed. Notice that for any catalytic reaction system $\mathcal{Q}$, any set $\mathcal{R}_i$ occurring in a CAF for $\mathcal{Q}$ is also a RAF for $\mathcal{Q}$.

Figure 1 illustrates these two concepts in the case of simple catalysation (of the form $c = c_C$).

There is a further condition we can impose on a RAF or CAF to make it more biologically relevant - namely we may require that a set of reactions is capable of constructing complex molecules required for maintaining certain biological processes (such as metabolism, error correction or reproduction). Of course there may be many combinations of complex molecules that suffice to maintain these processes, but we would like $\mathcal{R}'$ to be able to construct at least one of these combinations. We can formalize this notion as follows. Suppose $\mathcal{R}'$ is a RAF (respectively

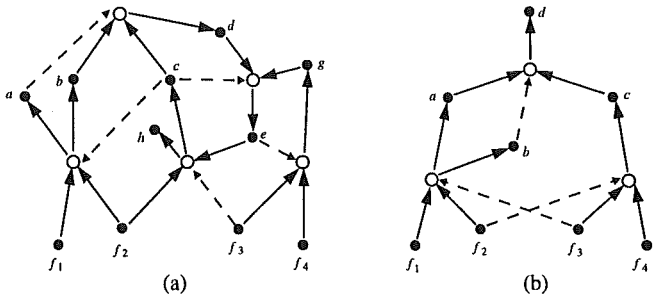

FIGURE 1. (a) An example of a RAF and (b) a CAF; represented as directed graphs. Molecules are shown as black nodes, reactions are white nodes, $F = \{f_1, f_2, f_3, f_4\}$ and each (positive) catalysation of a reaction by a molecule is indicated by dashed arc. Solid arcs show the input and output of each reaction.

$R_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_k = \mathcal{R}'$ a CAF) and suppose $\Omega \subseteq 2^{X-F}$ is a distinguished collection of subsets of molecules. We say that $\mathcal{R}'$ is an *(Ω)-complex RAF* (respectively an *(Ω)-complex CAF*) if the following condition (ΩC) holds:

(ΩC) If $\Omega \neq \emptyset$ then $C \subseteq \pi(\mathcal{R}')$ for at least one $C \in \Omega$.

We can think of each set $C \in \Omega$ as a suite of complex molecules that are required for maintaining certain biological processes and condition (ΩC) requires that the RAF or CAF be capable of constructing at least one such set. Note that the definition of an Ω–complex RAF (respectively Ω–complex CAF) reduces to that of a (simple) RAF or CAF if we take $\Omega = \emptyset$.

3. THE COMPLEXITY OF DETERMINING WHETHER OR NOT $\mathcal{Q}$ HAS A RAF OR A CAF

In Hordijk and Steel (2004) it was shown that, when $c = c_C$, there is a polynomial-time algorithm to determine if $\mathcal{Q}$ has a RAF. However if one allows inhibition also – by replacing c_C by $c = c_{C(+),C(-)}$– it is unlikely that any efficient algorithm exists for determining a RAF, by virtue of the following result.

Proposition 3.1. *For arbitrary catalytic reaction systems $\mathcal{Q} = (X, \mathcal{R}, c_{C(+),C(-)})$ and a subset F of X the decision problems 'does $\mathcal{Q}$ have a RAF?' is NP-complete.*

Proof. The decision problem is clearly in the class NP. To show it is NP-complete we provide a reduction from 3–SAT. Consider an expression P in conjunctive normal form involving binary variables $x_1, \ldots x_n$ and where each clause in P involves at most three variables. Thus we can write

$$P = C_1 \wedge C_2 \wedge \cdots \wedge C_k$$

where

$$C_i = \vee_{j \in T(i)} x_j \vee_{j \in F(i)} \overline{x}_j$$

and $T(i), F(i) \subseteq \{1, \ldots, n\}$, $|T(i)| + |F(i)| = 3$.

Given P construct a catalytic reaction system $\mathcal{Q} = (X, \mathcal{R}, c_{C(+),C(-)})$ as follows: let $F := \{x_1, \ldots, x_n\}$, let

$$X := \{x_1, \ldots, x_n, f_1, \ldots, f_n, t_1, \ldots, t_n, \theta_1, \ldots, \theta_k, 1\}.$$

Informally, x_i will correspond to the variable x_i in the formula; a reaction producing t_i (respectively f_i) will be catalyzed if the truth assignment of x_i is true (respectively false), and the reaction producing θ_i will be catalyzed if the i'th clause is satisfied.

More formally we let $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3$ where $\mathcal{R}_1$ consists of all reactions

$$x_i \xrightarrow{1(+),t_i(-)} f_i, \quad x_i \xrightarrow{1(+),f_i(-)} t_i,$$

for $1 \leq i \leq n$. In words x_i is the sole reactant for f_i and t_i but f_i inhibits the catalysation of t_i and vice-versa.

$\mathcal{R}_2$ consists of all reactions

$$\begin{cases} t_j \xrightarrow{1(+)} \theta_i & \text{if } j \in T(i), \\ f_j \xrightarrow{1(+)} \theta_i & \text{if } j \in F(i), \end{cases}$$

for $1 \leq j \leq n$ and $1 \leq i \leq k$. Finally $\mathcal{R}_3$ consists of the single reaction

$$\{\theta_1, \ldots, \theta_k\} \xrightarrow{1(+)} 1.$$

Now we claim that P has a satisfying truth assignment if and only if $\mathcal{Q}$ has a RAF. To establish this, first assume that P has a satisfying assignment. Fix such an assignment z and let $\{T, F\}$ be a partition of $\{1, \ldots, n\}$ corresponding to the variables that are true (respectively false) in z.

Now consider $\mathcal{R}_1' \cup \mathcal{R}_2' \cup \mathcal{R}_3$ where $\mathcal{R}_1' \subseteq \mathcal{R}_1$ consists of the reactions $x_i \to t_i$ for all $i \in T$, and the reactions $x_i \to f_i$ for all $i \in F$. $\mathcal{R}_2'$ will consist of the reactions $t_i \to \theta_j$ for all $i \in T \cap T(j)$ and $f_i \to \theta_j$ for all $i \in F \cap F(j)$. Since the assignment z satisfies the formula it follows that $\mathcal{R}_1' \cup \mathcal{R}_2' \cup \mathcal{R}_3$ is a RAF.

Next we have to show that if the system has a RAF the formula has a satisfying truth assignment. Suppose the system has a RAF $\mathcal{R}'$. Clearly $\mathcal{R}_3 \subset \mathcal{R}'$. This in turn implies that the reactions producing $\theta_1, \ldots, \theta_k$ are all catalyzed. Thus for all $1 \leq i \leq k$, there either exists some $j \in T(i)$ such that the reaction producing t_j is catalyzed or there exists some $j \in F(i)$ such that the reaction producing f_j is catalyzed. Moreover, for all i at most one of the reactions producing t_i and f_i can be catalyzed. We now define z_i to be true if the reaction producing t_i is catalyzed and false if the reaction producing f_i is catalyzed (z_i is defined arbitrarily otherwise). Then z is a satisfying assignment as required. □

However for any monotone catalysation function c there is a simple algorithm to determine whether or not $\mathcal{Q}$ has a CAF, which is essentially to let the system 'evolve from F'. We describe this now.

Proposition 3.2. *Given a catalytic reaction system $\mathcal{Q} = (X, \mathcal{R}, c)$, with c monotone, there is a polynomial time algorithm (in $|X|, |\mathcal{R}|$) to determine whether or not $\mathcal{Q}$ has a CAF.*

Proof. Define a sequence $X_i, \mathcal{R}_i$ for $i \geq 1$ as follows:

$$X_1 = F; \mathcal{R}_1 = \{r \in \mathcal{R} : \rho(r) \subseteq F, \text{ and } c(F, r) = 1\},$$

and for $i \geq 1$ set

$$X_{i+1} = X_i \cup \pi(R_i); R_{i+1} = R_i \cup \{r \in \mathcal{R} : \rho(r) \subseteq X_i \text{ and } c(X_i, r) = 1\}.$$

Then provided $\mathcal{R}_1 \neq \emptyset$ the sequence $\mathcal{R}_1 \subseteq \mathcal{R}_2 \cdots \subseteq \mathcal{R}_k$ (for any $k \geq 1$) is a CAF for $\mathcal{Q}$. If $\mathcal{R}_1$ is empty, then clearly $\mathcal{Q}$ has no CAF. □

4. Random sequence-based models

In this section we take $X = X(n)$, the set of sequences of length at most n over the alphabet set $\{0, 1, \ldots r-1\}$. Let F be a distinguished (small) subset of $X(n)$ - in this paper we will take $F = X(t)$ for a fixed value of t (often a value such as $t = 2$ has been taken in earlier papers). For a sequence $x \in X(n)$ we will let $|x|$ denote its length. Let $\mathcal{R}(n)$ denote the set of all ordered pairs $r = (A, B)$ where, for some $a, b, c \in X$ for which $c = ab$ (= the concatenation of a and b) either $A = \{a, b\}$ and $B = \{ab\}$ - in which case we call r a *forward reaction* or $A = \{ab\}$ and $B = \{a, b\}$ - in which case we call r a *backward reaction.* We may think of the pair $r = (\{a, b\}, c)$ as representing the ligation reaction

$$a + b \to c;$$

and the pair $r = (\{c\}, \{a, b\})$ as representing the cleavage reaction

$$c \to a + b.$$

We will let $\mathcal{R}_+(n)$ and $\mathcal{R}_-(n)$ denote the (partitioning) subsets of $\mathcal{R}(n)$ consisting of the forward and backward reactions, respectively.

Note that we have

$$x_n := |X(n)| = r + r^2 + \cdots + r^n = \frac{r^{n+1} - r}{r - 1}, \tag{2}$$

and

$$r_n := |\mathcal{R}_+(n)| = \left(r^2 + 2r^3 + ...(n-1)r^n\right) = \frac{(n-1)r^{n+2} - nr^{n+1} + r^2}{(r-1)^2}. \tag{3}$$

We will often below use the fact that, for all $n \geq 1$,

$$1 - O\left(\frac{1}{n}\right) \leq \frac{r_n}{nx_n} \leq 1 \tag{4}$$

(the notation $f(n) = g(n) + O(\frac{1}{n})$ means $|f(n) - g(n)| \leq K/n$ for some constant K for all $n \geq 1$).

We study a catalysation function c obtained by setting $c = c_C$ where C is some random assignment of catalysation (i.e. pairs (x, r)) that is subject to the following requirements:

(R1) The events $((x,r) \in C : x \in X(n), r \in \mathcal{R}_+(n))$ are independent.
(R2) For each sequence $x \in X(n)$ and reaction $r \in \mathcal{R}_+(n)$, the probability $\mathbb{P}[(x,r) \in C]$ depends only on x.

This model is more general than that described in Kauffman (1993), Steel (2000) or Hordijk and Steel (2004) for several reasons – it allows different catalysation probabilities for forward and backward reactions, it allows dependencies involving the catalysation of backward reactions, and the catalysation ability of a molecule can vary according to the molecule considered (for example, to depend on the length of the molecule).

Let $\mu_n(x)$ be the expected number of reactions in $\mathcal{R}_+(n)$ that molecule x catalyses. By (R2) we can write this as

$$\mu_n(x) = \mathbb{P}[(x,r) \in C] \cdot |\mathcal{R}_+(n)|,$$

for any given $r \in \mathcal{R}_+(n)$.

For $\mathcal{Q}(n) = (X(n), \mathcal{R}(n), c_C)$, $F = X(t)$ for some fixed value of t and $\Omega \subseteq 2^{X(n)-F}$, let $\mathcal{P}_n(\Omega)$ be the probability that $\mathcal{Q}(n)$ has an Ω–complex RAF. We can now state the first main result of this paper.

Theorem 4.1. *Consider a random catalytic reaction system $\mathcal{Q}(n)$ satisfying* (R1) *and* (R2) *and with $F = X(t)$ for a fixed value of t. Let $c \geq 0$ and let $\Omega \subseteq 2^{X(n)-F}$.*

(i) *Suppose that $\mu_n(x) \leq cn$ for all $x \in X(n)$. Then*

$$\mathcal{P}_n(\Omega) \leq 1 - \exp(-2cx_t^2(1 + O(\frac{1}{n}))\ (\to 0 \text{ as } c \to 0),$$

where x_t is defined in (2).

(ii) *Suppose that $\mu_n(x) \geq cn$ for all $x \in X(n)$, or that $\mu_n(x) \geq c\theta_n|x|$ for all $x \in X(n)$, where $c > \log_e(r)$ and where $\theta_n = \frac{1}{r}(1 + \frac{nr^{n+1}}{r_n}) \sim 1$. Then,*

$$\mathcal{P}_n(\Omega) \geq 1 - \frac{r(re^{-c})^t}{1 - re^{-c}}\ (\to 1 \text{ as } c \to \infty).$$

As an immediate corollary we obtain the following result, which confirms the two conjectures posed in Steel (2000).

Corollary 4.2. *Consider random catalytic reaction systems $\mathcal{Q}(n)$ $(n \geq t)$ satisfying* (R1) *and* (R2) *and with $F = X(t)$ for a fixed value of t. Take $\Omega = \emptyset$, and let $\mathcal{P}_n = \mathcal{P}_n(\emptyset)$, the probability that $\mathcal{Q}(n)$ has a RAF.*

(i) *If*

$$\max_{x \in X(n)} \frac{\mu_n(x)}{n} \to 0 \text{ as } n \to \infty$$

then $\lim_{n\to\infty} \mathcal{P}_n = 0$.

(ii) *If*

$$\min_{x \in X(n)} \frac{\mu_n(x)}{|x|} \to \infty \text{ as } n \to \infty$$

then $\lim_{n\to\infty} \mathcal{P}_n = 1$.

Remarks

- Corollary 4.2 has been worded in such a way that it clearly remains true if we interchange the terms $\frac{\mu_n(x)}{|x|}$ and $\frac{\mu_n(x)}{n}$ in either part (i) or part (ii) or both.
- The condition described in Corollary 4.2(ii) suffices to guarantee (for large n) a RAF involving all the molecules in $X(n)$. However it does not guarantee that all of $\mathcal{R}_+(n)$ is an *RAF*. The condition for this latter event to hold with high probability as $n \to \infty$ (assuming for simplicity that $\mu(x)$ is constant, say μ_n, over $X(n)$) is the stronger condition that

$$\liminf_{n\to\infty} \frac{\mu_n}{n^2} > \log_e(r).$$

  This follows from (a slight extension of) Theorem 1 of Steel (2000).
- Note that if we were to view a sequence $(x_1, x_2, \ldots x_n) \in X(n)$ and its reversal $(x_n, x_{n-1}, \ldots, x_1)$ as equivalent molecules then Corollary 4.2 still holds since asymptotically (with n) palindromic sequences have a negligible influence in the calculations.
- If we take $r = t = 2$ (the default setting for the simulations in Hordijk and Steel (2004)) then the lower bound on $\mathcal{P}_n$ provided in Theorem 4.1(ii) is positive for $c > 2\log(2)$. When $c = 4$ we have $\mathcal{P}_n > 0.99$.

To establish Theorem 4.1 we require first two further results - Lemma 4.3 and Proposition 4.4, and to describe them we introduce a further definition.

We say that a reaction $r \in \mathcal{R}(n)$ is *globally-catalyzed* (or GC) if there exists any molecule in $X(n)$ that catalyzes r. By the assumptions (R1) and (R2) above the probability that any forward reaction r is GC does not depend on r. Let p_* denote this probability and let $q_* = 1 - p_*$.

We will show that when p_* is sufficiently large then there exists a RAF $\mathcal{R} \subseteq \mathcal{R}_+(n)$ such that $X(n) - F \subseteq \pi(\mathcal{R})$ - in other words all molecules that are not already supplied by F can be generated by catalyzed reactions.

On the other hand, we will show that when p_* is small enough then the probability that there exists any globally catalyzed reaction that generates any molecule from $X(t+1)$ from molecules in $X(t)$ is small - thus proving that the probability that a RAF exists is small.

The first step is to estimate the probabilities of global catalysation.

Lemma 4.3. *Consider the system $\mathcal{Q}(n)$ satisfying properties* (R1) *and* (R2) *and with $F = X(t)$ for a fixed t, and let $c > 0$ be any positive constant.*

(i) *The probability q_* that a reaction $r \in \mathcal{R}_+(n)$ is not globally catalyzed is given by*

$$q_* = \prod_{x \in X(n)} \left(1 - \frac{\mu_n(x)}{r_n}\right).$$

In particular,

(ii) *if* $\mu_n(x) \leq cn$ *for all* x *then*

$$q_* \geq \exp(-c(1+O(\frac{1}{n})));$$

(iii) *if* $\mu_n(x) \geq cn$ *for all* x *then*

$$q_* < e^{-c}.$$

(iv) *if* $\mu_n(x) \geq c\theta_n|x|$ *for all* x *(where* θ_n *is as defined in Theorem 4.1) then*

$$q_* < e^{-c}.$$

Proof. Part (i) is immediate from (R1) and (R2). Part (ii) follows by combining part (i) and (4) to give:

$$q_* \geq \left(1-\frac{cn}{r_n}\right)^{x_n} \geq (1-\frac{c}{x_n(1-O(\frac{1}{n}))})^{x_n} = \exp(-c(1+O(\frac{1}{n}))).$$

Part (iii) follow from part (i) together with (4) which gives

$$q_* \leq \left(1-\frac{cn}{r_n}\right)^{x_n} \leq \left(1-\frac{c}{x_n}\right)^{x_n} < e^{-c},$$

as required. For part (iv), combine part (i), the identity $|\{x \in X(n) : |x| = s\}| = r^s$, and the inequality $(1-a)^b \leq \exp(-ab)$ for $a, b > 0$, to obtain

$$q_* \leq \prod_{s=1}^{n}\left(1-\frac{cs}{r_n}\right)^{r^s} \leq \prod_{s=1}^{n}\exp(-\frac{csr^s}{r_n}) = \exp(-\frac{c}{r_n}\sum_{s=1}^{n} sr^s).$$

Now, $\sum_{s=1}^{n} sr^s = r_{n+1}/r$ from (3), and part (iv) now follows by identifying θ_n with $\frac{r_{n+1}}{r_n r}$ (again using (3)). Note that θ_n converges to 1 as $n \to \infty$. □

Proposition 4.4. *Consider a random catalytic reaction system* $\mathcal{Q}(n)$ *satisfying properties* (R1) *and* (R2) *and with* $F = X(t)$ *for a fixed* t. *As before, denote the probability that a forward reaction is not globally catalyzed by* q_*. *Then*

(i) *The probability that* $\mathcal{Q}(n)$ *has a RAF is at most* $1 - q_*^{2x_t^2}$
(ii) *If* $rq_* < 1$ *then the probability that* $\mathcal{Q}(n)$ *has a RAF* $\mathcal{R}$ *with* $X(n) - F \subseteq \pi(\mathcal{R})$ *is at least*

$$1 - \frac{r(rq_*)^t}{1-rq_*}.$$

Proof. Part (i). Note that there are at most $2x_t^2$ forward reactions whose reactants (inputs) lie in $X(t)$. With probability $q_*^{2x_t^2}$ none of these reactions is GC, in which case there is no RAF for the system. The first part of the proposition now follows.

Part (ii). Note that, for any $s \geq t$ the probability that a molecule x with $|x| = s+1$ is not generated by any forward GC reaction from $X(s)$ is given by q_*^s. Therefore the expected number of molecules x with $|x| = s+1$ which are not generated by a forward GC reaction is $r^{s+1}q_*^s$. In particular the probability that there is a molecule in $X(s+1)$ that is not generated by a forward reaction from

$X(s)$ is at most $r^{s+1}q_*^s$. This in turn implies that the probability that all molecules in $X(n)$ are generated by forward GC reactions is at least

$$1 - r\sum_{s=t}^{n}(rq_*)^s \geq 1 - r\sum_{s=t}^{\infty}(rq_*)^s = 1 - \frac{r(rq_*)^t}{1-rq_*}.$$

Finally, note that if all molecules in $X(n) - F$ are generated by a set $\mathcal{R}$ of forward GC reactions, then $\mathcal{R}$ is a RAF for $\mathcal{Q}(n)$. □

Proof of Theorem 4.1

Part (i). By Proposition 4.4 (i) the probability that $\mathcal{Q}(n)$ has a RAF is at most $1 - q_*^{2x_t^2}$ which by Lemma 4.3(ii) is at most

$$1 - [\exp(-c(1 + O(\frac{1}{n})))]^{2x_t^2} = 1 - \exp(-2cx_t^2(1 + O(\frac{1}{n}))).$$

Clearly if $\mathcal{Q}(n)$ has no RAF, then it also has no Ω–complex RAF, for any $\Omega \subseteq 2^{X(n)-F}$.

Part (ii) This follows, by combining Proposition 4.4 (ii) with Lemma 4.3 parts (iii) and (iv), and noting that a RAF $\mathcal{R}$ of $\mathcal{Q}(n)$ for which $X(n) - F \subseteq \pi(\mathcal{R})$ is also an Ω–complex RAF for any $\Omega \subseteq 2^{X(n)-F}$. □

5. An analogous result for CAFs

The degree of catalysation required for a CAF to arise in the system $\mathcal{Q}(n)$ is much greater than for a RAF. This seems reasonable since the definition of a CAF involves a much stronger requirement than a RAF on a set of reactions. However the extent of the difference is interesting, and is given by the following analogue of Theorem 4.1.

Theorem 5.1. *Consider the random catalytic reaction system $\mathcal{Q}(n)$ and suppose that $F = X(t)$. Let $c \geq 0$ and let $\Omega \subseteq 2^{X(n)-F}$.*

(i) *If*

$$\mu_n(x) \leq \frac{c}{x_t^3} \cdot r_n$$

for all $x \in X(n)$, then the probability that $\mathcal{Q}(n)$ has a Ω-complex CAF is at most $2c$.

(ii) *If*

$$\mu_n(x) \geq \frac{c}{x_t} \cdot r_n,$$

for all $x \in X(n)$, then the probability that $\mathcal{Q}(n)$ has a Ω-complex CAF is at least

$$1 - \frac{r(re^{-c})^t}{1 - re^{-c}}.$$

Proof. Part (i). Let $\mathcal{R}' := \{r \in \mathcal{R}_+(n) : \rho(r) \subseteq F\}$, the set of all forward reactions that have all their reactants in F. The probability that any given reaction $r \in \mathcal{R}'$ is not catalyzed by at least one element of F is given by

$$\prod_{x \in F}(1 - \frac{\mu_n(x)}{r_n}).$$

Thus the probability that none of the reactions in $\mathcal{R}'$ are catalyzed by at least one element of F is

$$(\prod_{x \in F}(1 - \frac{\mu_n(x)}{r_n}))^{|\mathcal{R}'|}.$$

In particular if $\mu_n(x) \leq \frac{cr_n}{x_t^3}$, then, since $|F| = x_t$ and $|\mathcal{R}'| = 2x_t^2$, this probability (that none of the reactions in $\mathcal{R}'$ is catalyzed by at least one element of F) is at least

$$(1 - \frac{c}{x_t^3})^{2x_t^3} \geq 1 - 2c.$$

However when none of the reactions in $\mathcal{R}'$ is catalyzed, then $\mathcal{Q}(n)$ does not have a CAF. Thus the probability that $\mathcal{Q}(n)$ has a CAF is at most $2c$.

Part (ii). For every molecule in $x \in X(n)$, and each $s \in \{t, \ldots, n\}$ let $E_s(x)$ be the event that there is at least one reaction r_x of the form $a + b \to x$, where $a, b \in X(s)$, that is catalysed by at least one molecule in $X(s)$.

Now, if $\mu_n(x) \geq \frac{cr_n}{x_t}$, then for any forward reaction r, the probability that r is not catalysed by at least one molecule in $X(s)$ (for $s \geq t$) is at most

$$(1 - \frac{c}{x_t})^{x_s} \leq \exp(-c\frac{x_s}{x_t}) \leq e^{-c}$$

and since, for each x there are $|x| - 1$ choices for r_x we have

$$\mathbb{P}(E_s(x)^c) \leq \exp(-c(|x| - 1)), \tag{5}$$

where $E_s(x)^c$ is the complementary event to $E_s(x)$.

Consider the event

$$E_s := \bigcap_{x \in X(s+1) \setminus X(s)} E_s(x).$$

By (5) and the identity $|X(s+1) - X(s)| = r^{s+1}$ we have $P(E_s^c) < r^{s+1}e^{-cs}$ and so

$$\mathbb{P}(\bigcap_{s=t}^{n-t} E_s) \geq 1 - \sum_{s=t}^{\infty} r^{s+1}e^{-cs} = 1 - \frac{r(re^{-c})^t}{1 - re^{-c}}.$$

However the event $\bigcap_{s=t}^{n-t} E_s$ ensures that the nested collection of reactions $\mathcal{R}_i := \{r_x : x \in X(t+i)\}$, $i = 1, \ldots, n-t$ forms a CAF for $\mathcal{Q}(n)$, and moreover one for which the maximal set $\mathcal{R}_{n-t}$ generates all elements of $X(n) - F$ - thus it is also a Ω–complex CAF for any $\Omega \subseteq 2^{X(n)-F}$. This completes the proof.

□

6. Conclusion

The question of how life first arose on earth is a multifaceted problem that stands out as one of the major questions in science (see for example Dyson, 1985; Fenchel, 2002; Joyce, 1989). The theoretical investigation of catalytic reaction systems is an attempt to address just one aspect of this question. This concerns the issue of whether, as Kauffman has maintained, we should expect self-sustaining, autocatalytic networks to emerge in random chemical systems once some threshold (in 'complexity', 'connectivity' or 'catalysation rate') is exceeded, or whether there is the requirement of some fine-tuning of the underlying biochemistry for such networks to occur. Our results here have helped delineate precisely how much catalysation is required in order for random sequence-based chemical reaction systems (without any 'fine-tuning') to likely give rise to a RAF. In contrast to a CAF, where a very high degree of catalysation is required when the maximal sequence length n is large, the likely occurrence of a RAF depends just on whether the catalysation function $\mu_n(x)$ grows sublinearly or superlinearly with n (Corollary 4.2). One question for future work would be to explore how the results in this paper would be influenced by allowing random inhibitory catalysations. The model studied here could also be refined to better suit the graph-theoretic properties of real metabolic networks which have recently been investigated (Jeong *et al.*, 2000; Wagner and Fell, 2001).

7. Acknowledgements

The first author is supported by a Miller fellowship. The second author thanks the NZIMA (Maclauren Fellowship) for supporting this work.

REFERENCES

[1] Bollobas, B., 2001. Random graphs, 2nd Edition. Cambridge Studies in Advanced Mathematics, Vol. 73. Cambridge University Press, Cambridge.

[2] Bollobas, B., Rasmussen, S., 1989. First cycles in random directed graph processes. Discr. Math. 75, 55-68.

[3] Cohen, J.E., 1988. Threshold phenomena in random structures. Discr. Appl. Math. 19, 113–128.

[4] Dyson, F., 1985. Origins of Life. Cambridge University Press, Cambridge.

[5] Eigen, M., 1971. Self-organisation of mater and evolution of biological macromolecules, Naturwissenshaften 58, 465–523.

[6] Eigen, M., Schuster, P., 1979. The hypercycle - a principle of natural selforganisation (Springer-Verlag, Heidelberg.

[7] Fenchel, T., 2002. The origin and early evolution of life. Oxford University Press, Oxford.

[8] Hordijk, W., Fontanari, J.F., 2003. Catalytic reaction sets, decay, and the preservation of information. Proceedings of KIMAS03 (IEEE International Conference on Integration of Knowledge Intensive Multi-Agent Systems, Boston, MA). pp. 133–138.

[9] Hordijk, W., Steel, M., 2004. Detecting autocatalyctic, self-sustaining sets in chemical reaction systems *J. Theor. Biol.* 227(4): 451-461.

[10] Jeong, H., Tombor, B, Albert, R. Oltvai, Z.N. and Barabási. (2000), The large-scale organisatoin of metabolic networks. Nature 407, 651–654.

[11] Joyce, G.F., 1989. RNA evolution and the origin of life. Nature 338, 217–224.

[12] Kauffman, S.A., 1986 Autocatalytic sets of proteins, J. Theor. Biol. 119, 1–24.

[13] Kauffman, S.A., 1993. The origins of order: self-organisation and selection in evolution. Oxford University Press, Oxford.

[14] Lifson, S., 1997. On the crucial stages in the origin of animate matter. J. Mol. Evol., 44, 1–8.

[15] Lohn, J.D., Comombano, S.P., Scargle, J., Stassinopoulos, D., Haith, G.L., 1998. Evolving catalytic reaction sets using genetic algorithms. Proceedings of the 1998 IEEE International Conference on Evolutionary Computation, Anchorage, AK, USA. pp. 487–492.

[16] Maynard Smith. J., 1983. Models of evolution. Proc. Roy. Soc. Lond. B 219: 315–325.

[17] Scheuring, I., 2000. Avoiding catch-22 by early evolution of stepwise increase in copying fidelity. Selection 1: 135–145.

[18] Steel, M., 2000. The emergence of a self-catalysing structure in abstract origin-of-life models, Appl. Math. Lett. 3, 91-95.

[19] Wagner, A. and Fell, D.A., 2001. The small world inside large metabolic networks. Proc. R. Soc. London B, 268, 1803–1810.

[20] Wills, P. and Henderson, L., 1997. Self-organisation and information-carrying capacity of collectively autocatalytic sets of polymers: ligations systems in: Y. Bar-Yam (Ed.) Proceedings of the International Conference on Complex Systems, New England Complex Systems Institute.

[21] Zintzaras, E., Santos, M. and Szathmáry, E., 2002. "Living" under the challenge of information decay: The stochastic corrector model vs hypercycles. J. Theor. Biol. 217, 167–181.

ELCHANAN MOSSEL: STATISTICS DEPARTMENT, UC BERKELEY, CA, USA; MIKE STEEL: ALLAN WILSON CENTRE FOR MOLECULAR ECOLOGY AND EVOLUTION, UNIVERSITY OF CANTERBURY, CHRISTCHURCH, NEW ZEALAND

E-mail address: mossel@stat.berkeley.edu, m.steel@math.canterbury.ac.nz